

OSVRT -

Boja i zvuk u video kompresiji

Ovim predavanjem se osvrćemo na temu kompresije podataka koja je važan čimbenik u prijenosu video i zvučnog signala. Prijenos podataka, odnosno bit-rate, je u rasponu od 270 megabita po sekundi za SDTV, 1.5 gigabita za hdtv ili 3 ili više gigabita po sekundi za UHDTV. Takva količina podataka je zapravo prevelika pa svi podaci moraju biti komprimirani koristeći perceptualno kodiranje da bismo smanjili količinu podataka koje se šalje. Perceptualno kodiranje uzima u obzir karakteristike ljudskih organa, a to su naše oči i uši. Te operacije se izvode prilikom transmisije video signala ili prilikom samog snimanja kamerama, fotoaparatom itd. . Svi oni imaju mali kapacitet pohrane i moraju izvoditi neku vrstu kompresije da bi mogli spremiti podatke. Također, kod transmisije, moramo voditi računa o količini podataka koju šaljemo kako bi primatelj signala mogao neometano pratiti signal bez prevelikih gubitaka. Kad govorimo o sažimanju podataka direktnom prilikom snimanja, govorimo o kodiranju izvora, odnosno source-coding jer sažimamo analogni signal koji dolazi na senzore i uređaja za snimanje. Za razliku od komprimiranja podataka u post procesiranju, kako bi prilagodio za različite medije , razlikujemo 2 vrste kodiranja odmah na izvoru i kasnije u postprocesiranju. Nakon kodiranja izvora, količina podataka će biti svedena na 1-15 megabita ovisno o algoritmu kompresije koju koristimo.

Postoji reduciranje suvišnih i nevažnih podataka. Suvišni (redundantni) podaci su oni koji se ponavljaju više puta, a nevažni (irelevantni) su oni koje ljudsko oko ne će primijetiti ukoliko nedostaju. Oni mogu smanjiti količinu podataka i preko 100 puta. Redundantni podaci mogu se lako i bez gubitaka izračunati nekim matematičkim algoritmima prilikom dekodiranja (*lossless* kompresija). Kompresija irelevantnih podataka naziva se kompresija s gubitcima (*lossy* kompresija).

- Redukcija podataka za boju
- RGB -> luminantna i krominantna komponenta
- YCbCr (YUV)

$$Y = (0.3 * R) + (0.59 * G) + (0.11 * B)$$

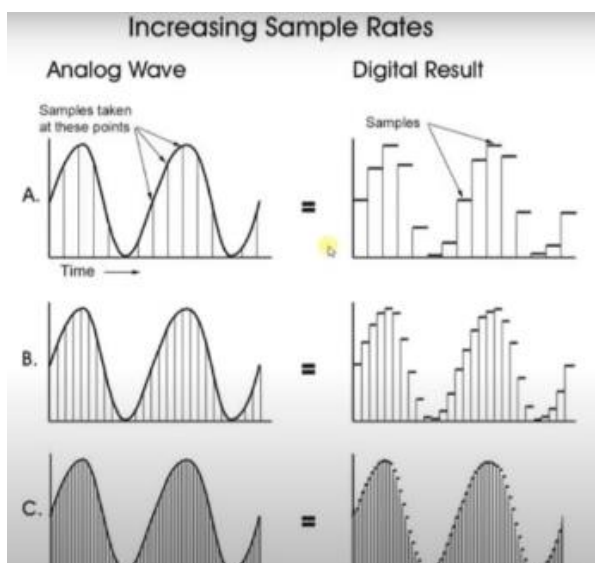
$$Cb = 0.56 * (B - Y)$$

$$Cr = 0.71 * (R - Y)$$

Ljudsko uho ima dinamički raspon od oko 140 decibela i raspon frekvencija od 20 herca do 20 000 herca. Valovi veće frekvencije od 20 000 herca nazivaju se ultrazvuk, a frekvencije manje od 20 herca nazivaju se infrazvuk. Audio signal koji želimo transmitirati digitalnim putem mora imati karakteristike koje poštuju raspone čujnosti ljudskog uha.

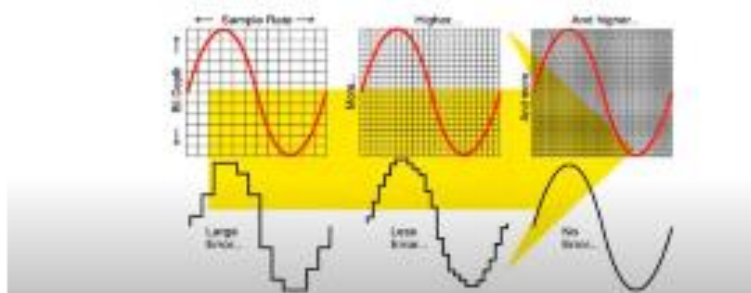
Analogni signal snima se različitim uređajima te se putem filtera limitira njegov raspon prije nego što se digitalizira. Taj proces zove se *sempliranje* ili *uzorkovanje* signala. Time analogni signal činimo mjerljivim.

Sampling rate ili brzina uzorkovanja mjeri se kao broj snimljenih audio uzoraka unutar jedne sekunde.



Sampling depth ili preciznost uzorka mjeri se u bitovima po uzorku. Ta mjera određuje koliko mogućih razina amplitude može postići zvučni signal.

Sampling depth – preciznost uzorka



$$\text{Bit rate} = \text{Sample rate} * \text{Sample depth}$$

$$= \text{uzorak / sek} * \text{bit / uzorku} = \text{bit / sek}$$

Kompresija podataka => redundantni i irelevantni podaci

- Najčešće korišteni sampling rate:
- 8 kHz – telefonija
- 44.1/48 kHz – TV / CD
- 96/192 kHz – blu-ray
- > 300 kHz

